Jet Propulsion Laboratory
California Institute of Technology

PDS4 Development How-To: Labels, Bundles & Local Data Dictionaries

Planetary Data Workshop

Flagstaff, AZ -- June 19, 2019

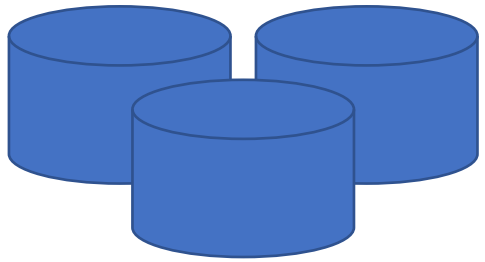
Cristina M. De Cesare
cristina.m.decesare@jpl.nasa.gov
Mission Interface Lead, PDS Cartography and Imaging Sciences Node
Jet Propulsion Laboratory, California Institute of Technology

Outline

- What is PDS? What is PDS4?
- How do I develop PDS4 labels?
- What are PDS4 dictionaries & which ones should I use?
- How do I develop PDS4 bundles?

What are PDS and PDS4?

The **Planetary Data System** (PDS) is NASA's repository for the distribution and long term preservation of NASA planetary data.



The **PDS Archive** is the digital data repository maintained by PDS.

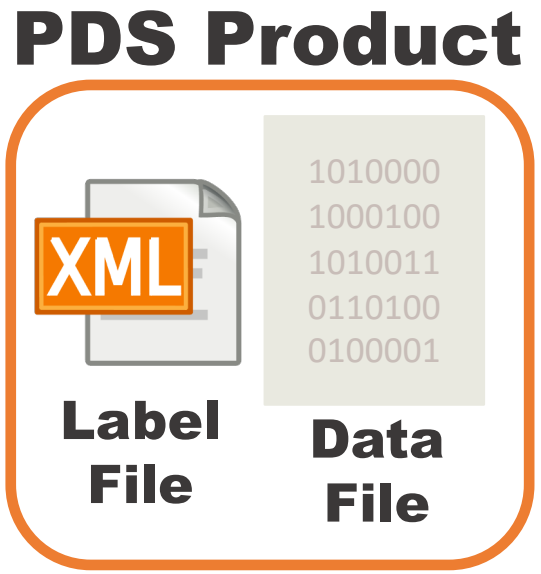
The **PDS Standard** are requirements and constraints designed to ensure the usability of data in the PDS Archive throughout the lifetime of the archive.

PDS4 is the latest version of the PDS Standard. (PDS4 is **not** a data format.)

Outline

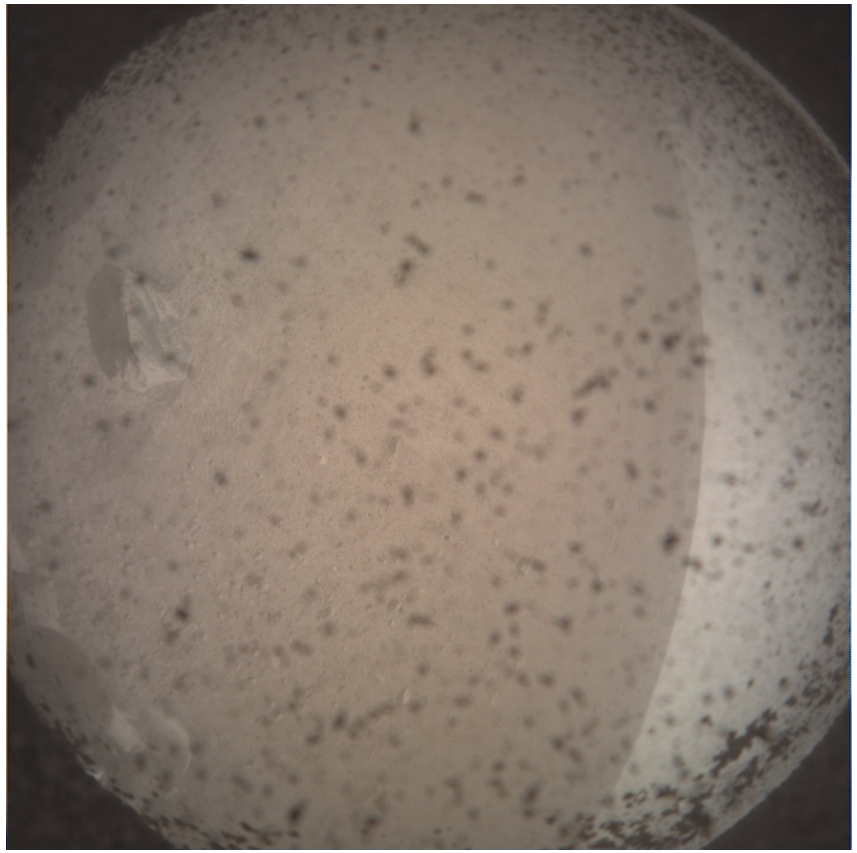
- What is PDS? What is PDS4?
- How do I develop PDS4 labels?
- What are PDS4 dictionaries & which ones should I use?
- How do I develop PDS4 bundles?

What are PDS4 labels?



- **PDS4 labels** are XML files that describe the contents of one or more science data files:
  - Observational data
  - Documentation files
  - Calibration data
- A PDS label, along with the file(s) that it describes, constitute a **PDS Product**.
- Every file in a PDS4 archive must be described in a **PDS4 label**.

InSight case study: example label metadata

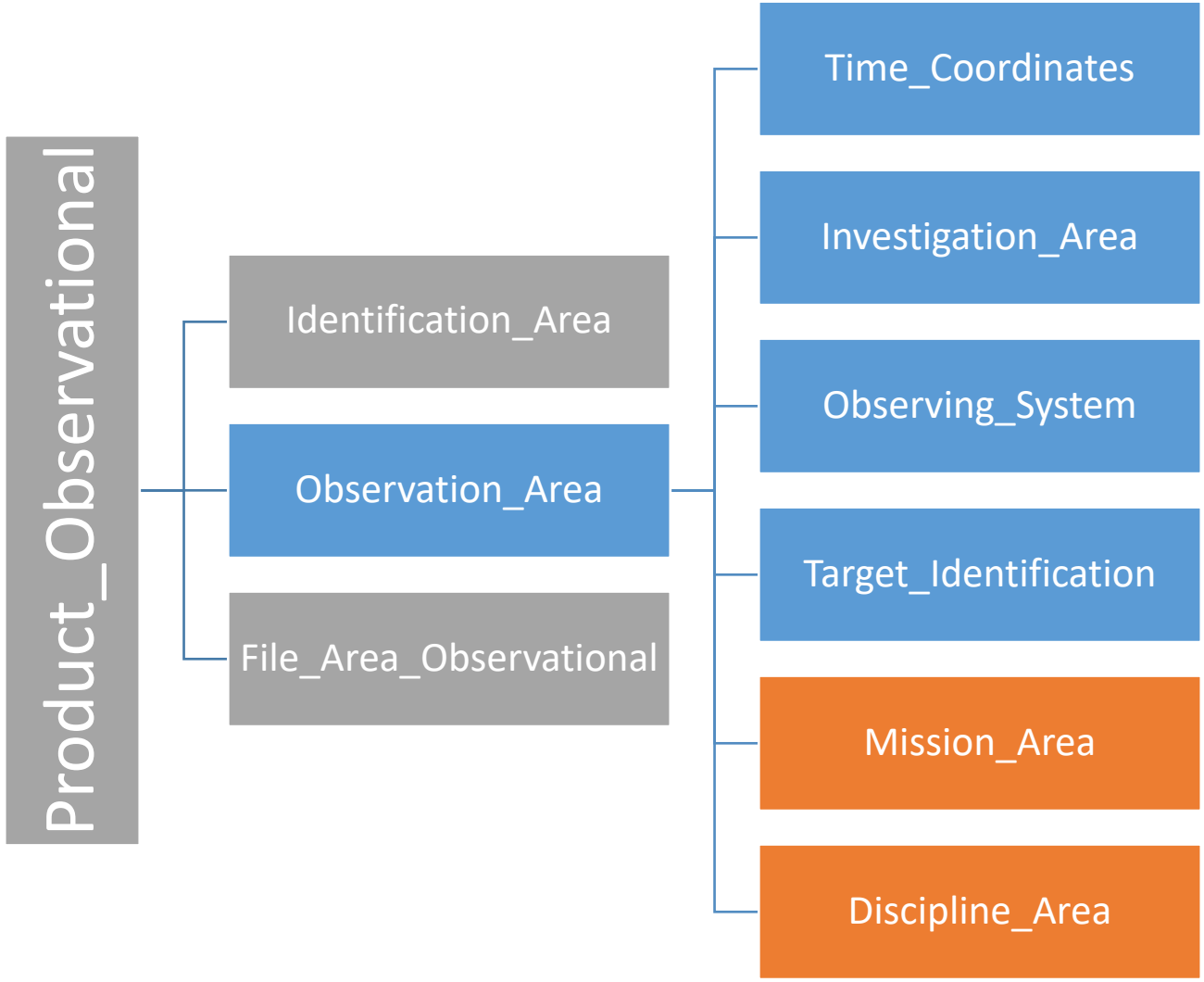


Observational data file:
raw image from InSight sol 0

```
<img:Exposure>
   <img:exposure_count>2</img:exposure_count>
   <img:exposure_duration unit="ms">329.13</img:exposure_duration>
   <img:exposure_duration_count>53</img:exposure_duration_count>
</img:Exposure>
```

```
<msn:mission_phase_name>SURFACE MISSION</msn:mission_phase_name>
<msn:product_type_name>EDR</msn:product_type_name>
<msn:spacecraft_clock_start>596533559.23421</msn:spacecraft_clock_start>
<msn:spacecraft_clock_stop>596533559.44990</msn:spacecraft_clock_stop>
```

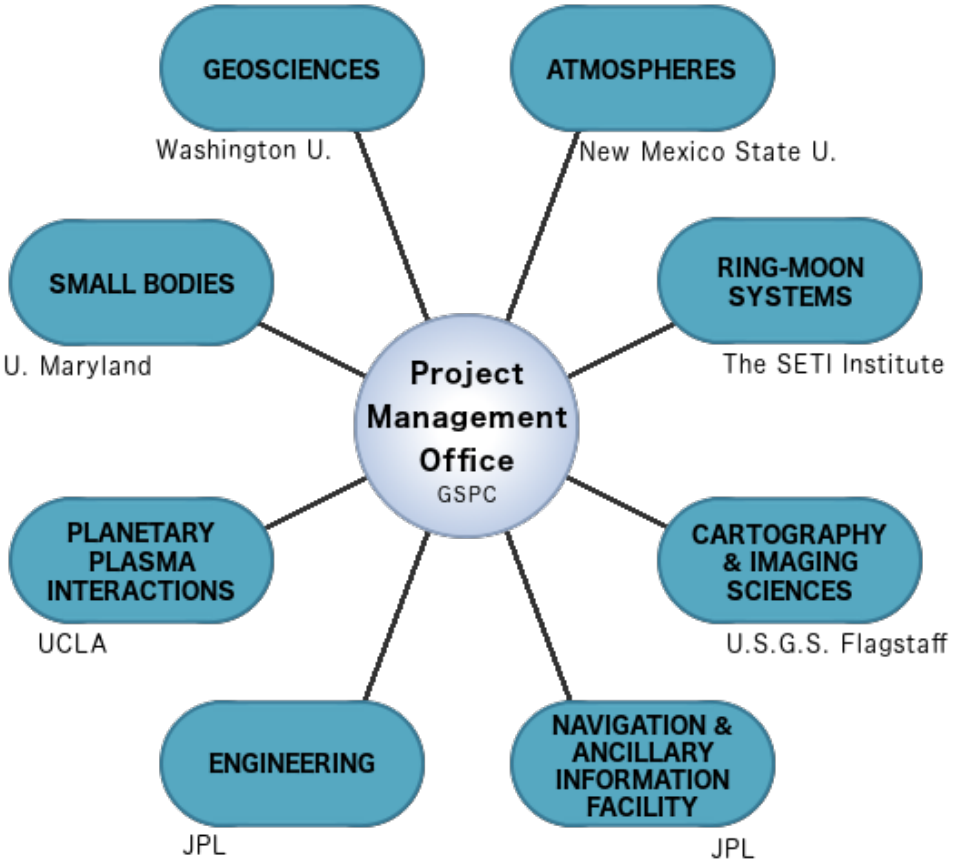
What does a PDS4 label look like?



I need PDS4 labels. Where do I start?

Ask yourself these questions:

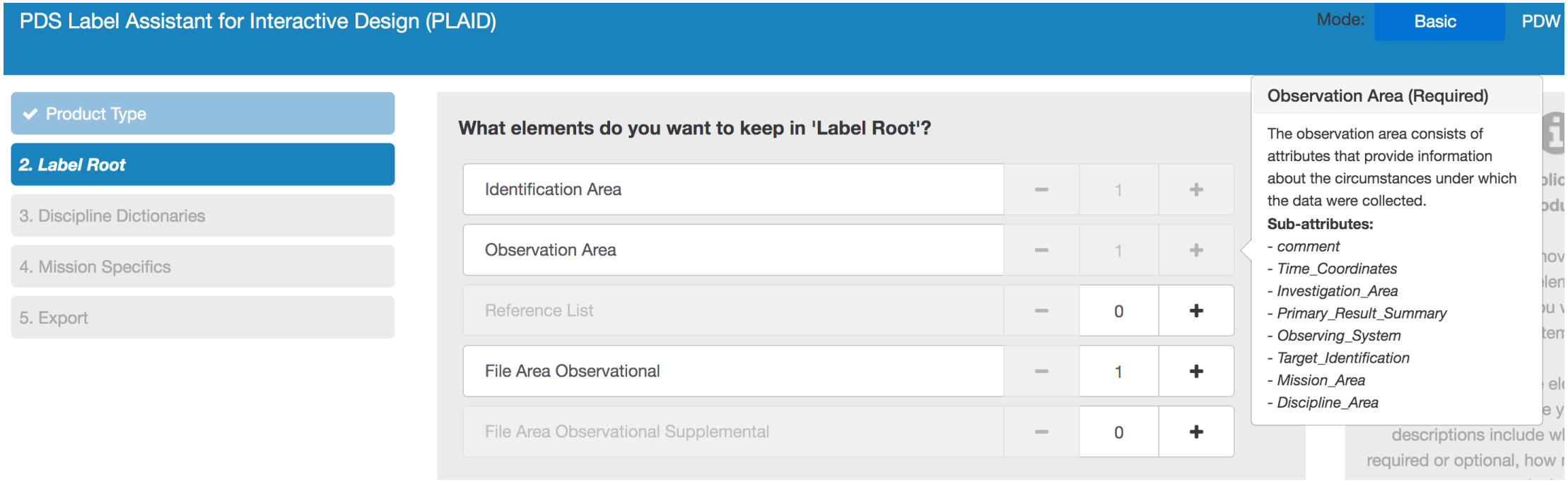
- What kinds of data products will I be archiving?
- Are there many different types of data products?
  - Does each product type have unique metadata?
- Which PDS Node will this data be archived with?
  - If unsure, contact pds_operator@jpl.nasa.gov



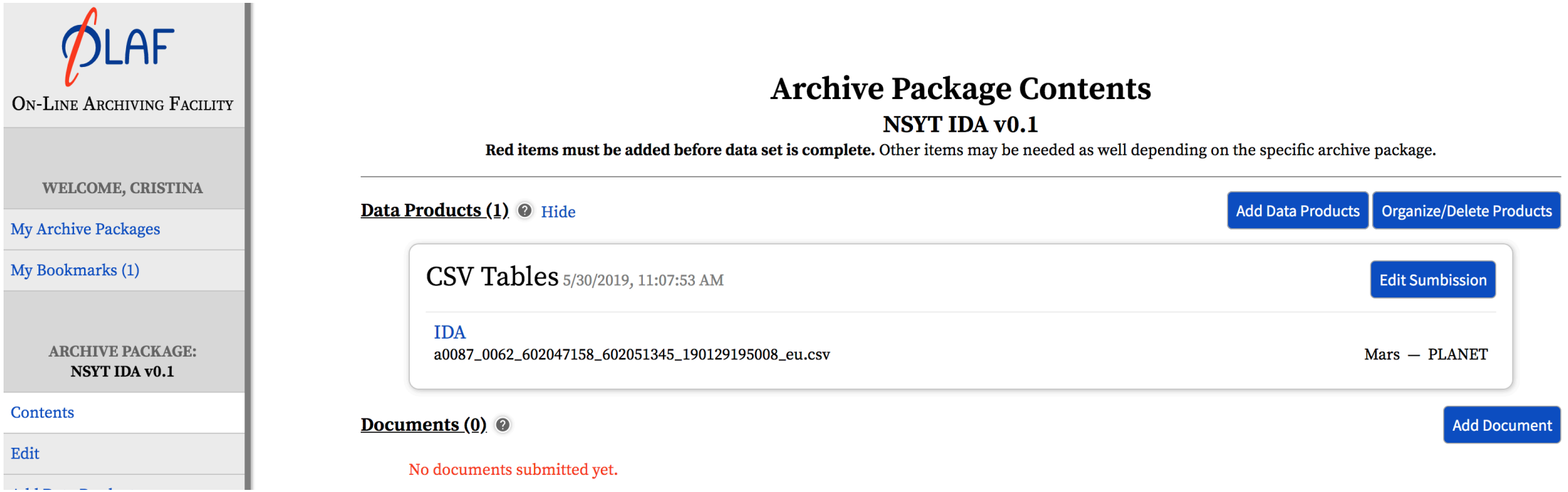
How do I create my 1st label?

1. Choose 1 data file to start with.
2. Install a text editor or XML editor
    1. Notepad++
    2. OxygenXML
3. Decide which PDS tool(s) to use:
    - **PLAID** (PDS Label Assistant for Interactive Design) is a good starting point. It's an interactive web tool that will help you build a PDS4 label step-by-step.
    - Is your data file a CSV or fixed-width table?
        - Try **OLAF (On-Line Archiving Facility)**, a web tool that will help you build a complete PDS4 archive bundle from the data files you upload.

PDS Label Assistant for Interactive Design (PLAID)



On-Line Archive Facility (OLAF)



How do I validate my PDS4 label?

- Use PDS4 Validate Tool
- If using Oxygen XML Editor (or other XML editor), you can also validate your label from inside the software tool.

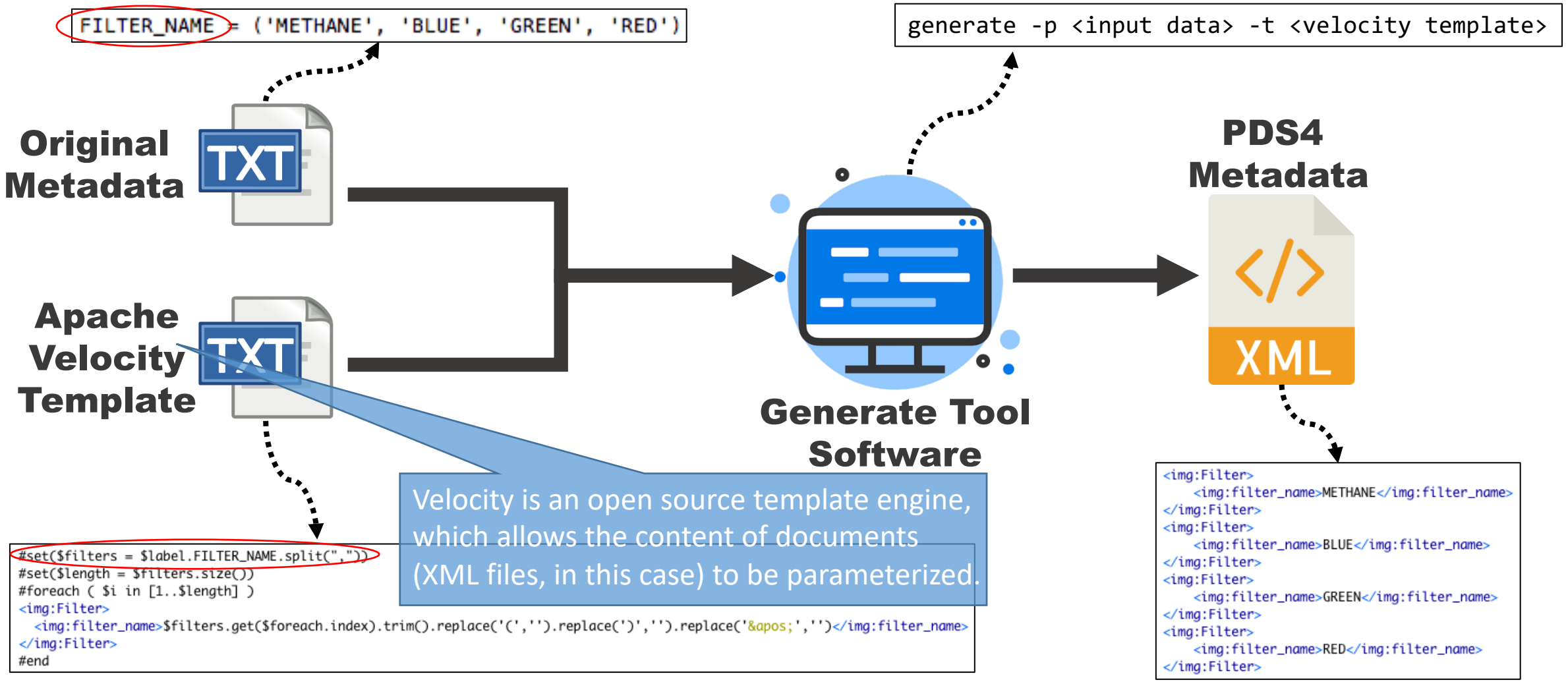
What if my label doesn't validate?

- Don't panic!
- The validation report will show you a list of errors.
  - Don't forget, in PDS4, ordering matters.
- If you're stuck, contact your PDS Node representative for help.

I have 1 valid label. Now what?

- Use this prototype label as a model for the rest of your labels:
  1. Transform your prototype into a label template, by replacing specific metadata values with variables.
  2. Write a script to read in your science data files & replace the variables with real metadata.
- If your science data files include ODL metadata files or you're doing a migration from PDS3….
  - You can use the PDS Generate Tool to perform the above process.

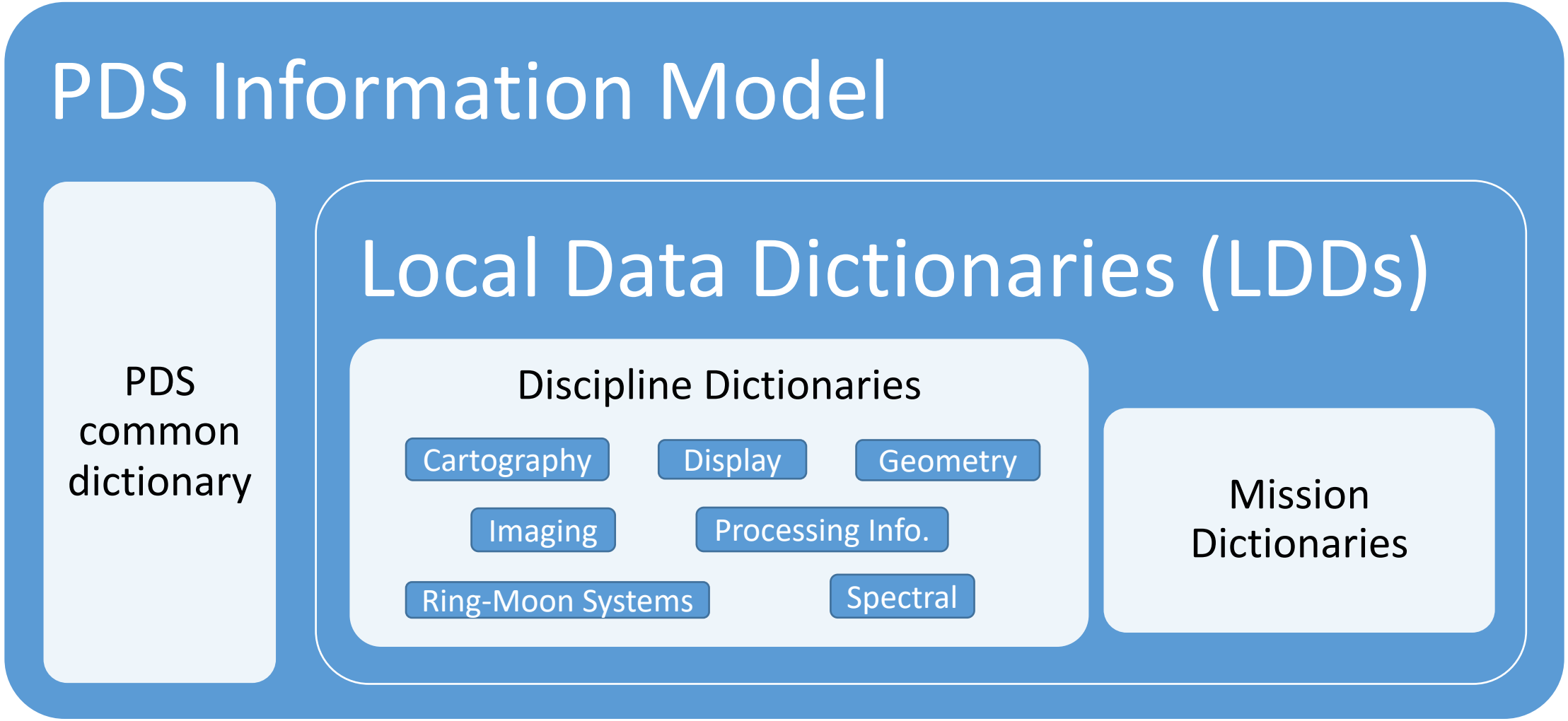
Using Generate to produce labels



Outline

- What is PDS? What is PDS4?
- How do I develop PDS4 labels?
- What are PDS4 dictionaries & which ones should I use?
- How do I develop PDS4 bundles?

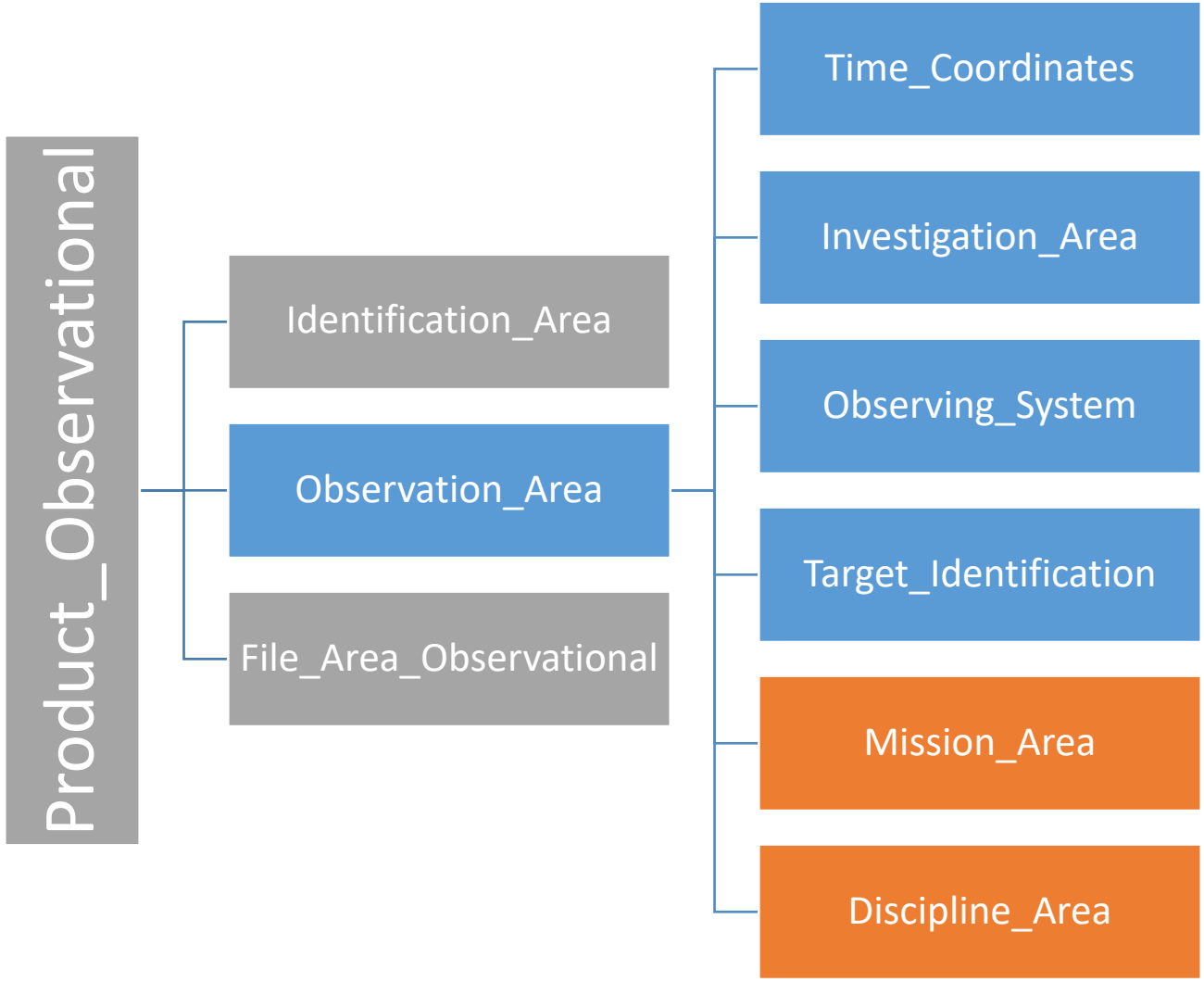
3 categories of data dictionary



What are PDS4 data dictionaries?

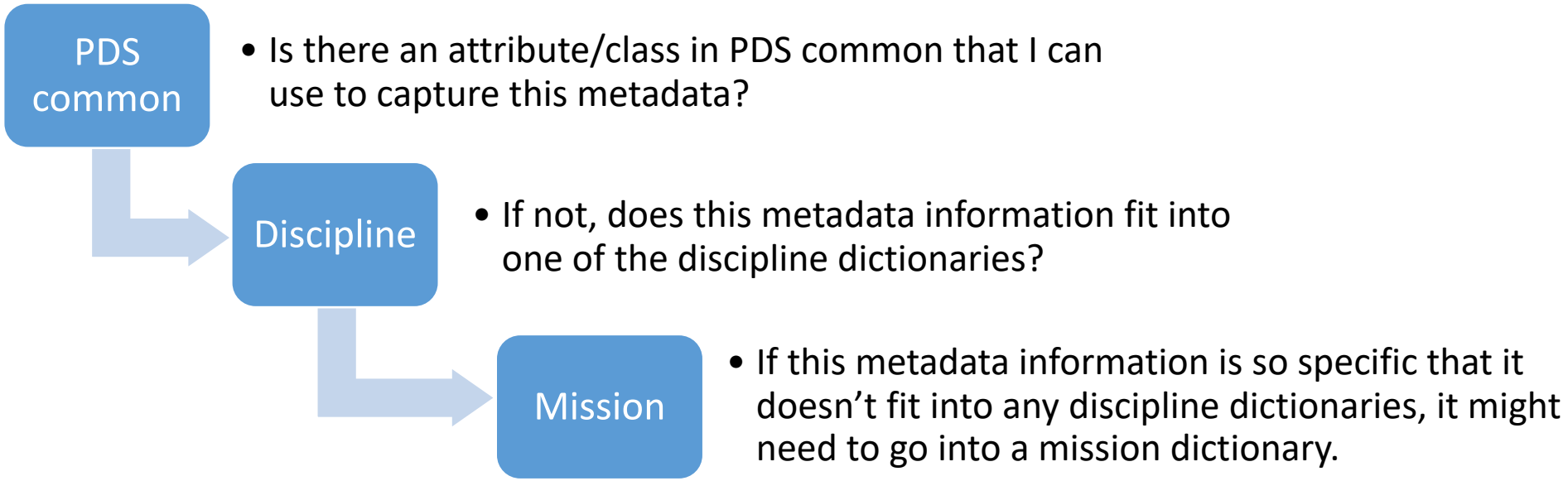
- PDS4 data dictionaries…
  - Define the meaning and structure of XML classes & attributes to be used in PDS4 label files.
  - Specify XML tags, their meanings, and acceptable values
  - Include an XML Schema file (*.xsd) & XML Schematron file (*.sch)
- 3 categories of dictionary:
  - PDS core ← Least specific
  - Discipline dictionary
  - Mission dictionary ← Most specific
  - The latest PDS4 dictionaries are published on the PDS website:
    - https://pds.nasa.gov/datastandards/schema/released/

Remember this?



Which dictionaries should I use?

- Every PDS4 label must use the PDS common dictionary.
  - It defines the essential, required components of even the simplest label.
- Discipline & mission dictionaries are optional, and their use will be driven by the types of metadata you wish to include in your labels.



What if my metadata doesn't fit into any dictionary?

- Ask your PDS Node representative for help.
- They can advise you on…
  - Whether any existing data dictionaries could be updated in order to support your metadata.
    - And what this process looks like
  - Whether you ought to create your own mission-specific local data dictionary
    - And how to do so.
    - Refer to Anne Raugh's upcoming talk about building Local Data Dictionaries!

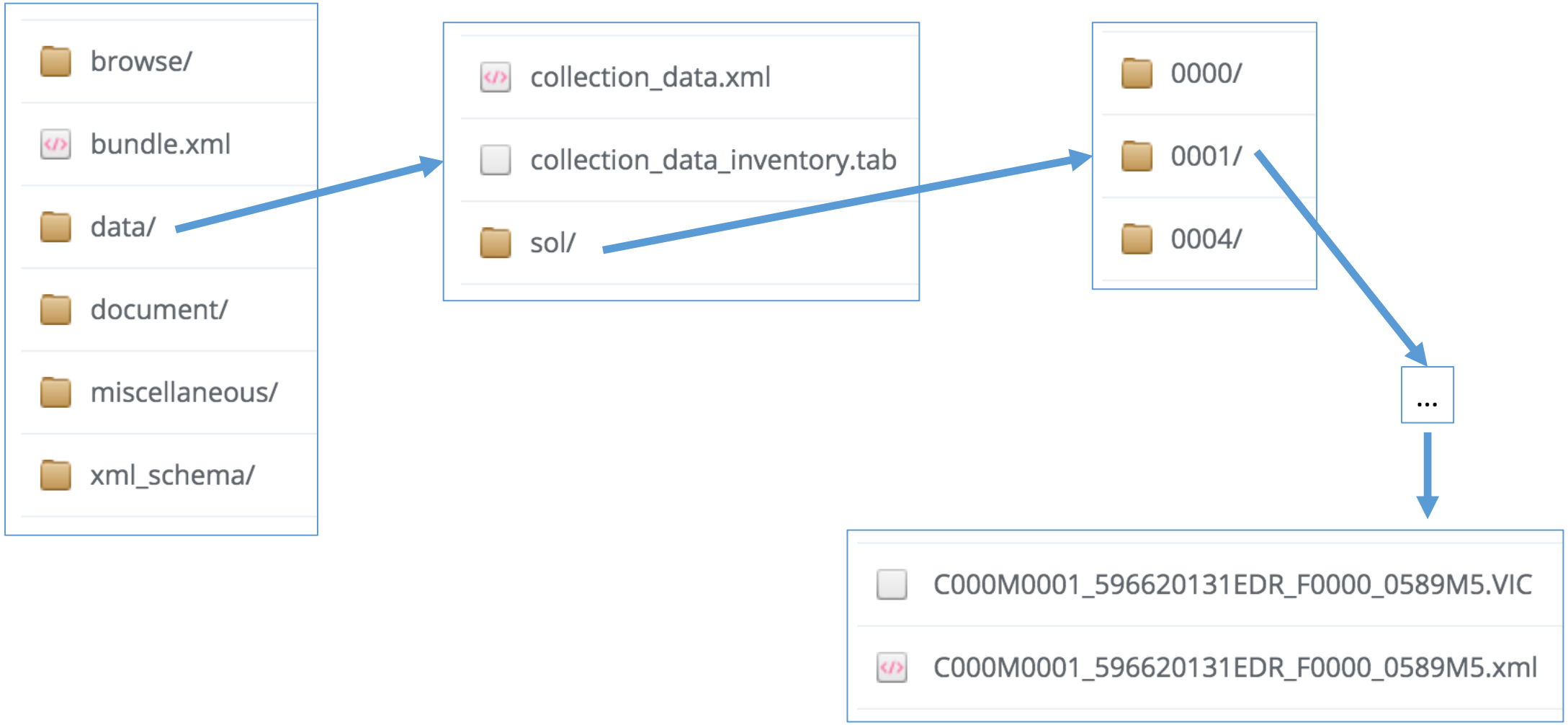
Outline

- What is PDS? What is PDS4?
- How do I develop PDS4 labels?
- What are PDS4 dictionaries & which ones should I use?
- How do I develop PDS4 bundles?

I’ve got my labels & dictionaries. How do I put them all together?

- Collections
  - Products are grouped into ‘collections’ by type and content.
  - Examples:
    - Data collection
    - Document collection
    - Supplementary collection
- Bundle
  - A list of collections
  - Some missions may have 1 bundle, while others may have multiple
  - This is what you will package up & deliver to your curating Node!

InSight case study: archive design



How do I develop PDS4 bundles?

- Write your own scripts to:
  - Generate labels into their respective collection directories
  - Place those collection directories into a bundle directory
  - Generate collection-level and bundle-level metadata files
  - Invoke the PDS4 Validate Tool on the new bundle
- Scripts can be integrated into existing data processing pipelines for a mission/observation.
- Or, consider using services provided by AMMOS (Advanced Multimission Operations System)
  - PDS Archive Transformation Service (PATS)
    - Provides a method to transform PDS3 archive to PDS4
  - PDS Ongoing Delivery Service (PODS)
    - Provides a method to process & deliver PDS4-compliant data

Resources

AMMOS (PATS & PODS)	ammos.nasa.gov
Data Provider's Handbook	https://pds.nasa.gov/datastandards/documents/dph/current
Generate Tool	https://pds.nasa.gov/tools/about/generate/
OLAF	https://olaf.psi.edu/
PDS Node Contacts	https://pds.nasa.gov/contact/contact.shtml
PDS4 Schemas	https://pds.nasa.gov/datastandards/schema/released/
PLAID	plaid.jpl.nasa.gov
Validate Tool	https://pds.nasa.gov/tools/about/validate/

References

1. NASA PDS. (2018) *What is PDS4?* Retrieved from https://pds.jpl.nasa.gov/pds4/about/what.shtml .
2. NASA PDS. (2018) *PDS Concepts*. Retrieved from https://pds.jpl.nasa.gov/datastandards/documents/concepts/Concepts_1.11.0.pdf
3. Moses Milazzo. (2018) *Introduction to PDS4*. Retrieved from https://pds.nasa.gov/datastandards/training/2018-ec/ec-pds4-introduction-201806.pdf
4. NASA PDS. (2017) *PDS4 Training Exercise*. Retrieved from https://pds.nasa.gov/datastandards/training/2017-agu/pds4-training-agu-2017-v2.pptx
5. Jet Propulsion Laboratory. (2018) *InSight ICC EDR Observational Product - c000m0000_596533559edr_f0000_0106m*. Retrieved from https://pds-imaging.jpl.nasa.gov/data/nsyt/insight_cameras/data/sol/0000/mipl/edr/icc/C000M0000_596533559EDR_F0000_0106M9.xml
6. "Text-xml" by RRZEicons is licensed under CC BY 2.0. https://commons.wikimedia.org/wiki/File:Text-xml.svg.
7. "Text-txt" by RRZEicons is licensed under CC BY 2.0. https://commons.wikimedia.org/wiki/File:Text-xml.svg.

Questions?

Jet Propulsion Laboratory
California Institute of Technology

jpl.nasa.gov